

INTERNATIONAL SEARCH REPORT

Int. Patent Application No
PCT/EP2004/051784

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G11B19/02

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G11B

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|------------|---|-----------------------|
| X | US 5 617 407 A (BAREIS MONICA M) 1 April 1997 (1997-04-01) the whole document | 1-14 |

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *&* document member of the same patent family

Date of the actual completion of the international search

5 January 2005

Date of mailing of the international search report

13/01/2005

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Benfield, A

Information on patent family members

PCI/EP2004/051784

US 5617407 A 01-04-1997 NONE

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP2004/051784

A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES
IPK 7 G11B19/02

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)
IPK 7 G11B

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

EPO-Internal, WPI Data, PAJ

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

| Kategorie* | Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile | Betr. Anspruch Nr. |
|------------|--|--------------------|
| X | US 5 617 407 A (BAREIS MONICA M) 1. April 1997 (1997-04-01) das ganze Dokument ----- | 1-14 |



Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen



Siehe Anhang Patentfamilie

* Besondere Kategorien von angegebenen Veröffentlichungen :

A Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist

E Älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

L Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)

O Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

P Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

T Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

X Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden

Y Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

Z Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

5. Januar 2005

Absenddatum des internationalen Recherchenberichts

13/01/2005

Name und Postanschrift der internationalen Recherchenbehörde
Europäisches Patentamt, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Benfield, A

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

PCT/EP2004/051784

| | | | |
|------------|---|------------|-------|
| US 5617407 | A | 01-04-1997 | KEINE |
|------------|---|------------|-------|

Beschreibung

Sprachsteuerung von Audio- und Videogeräten

5 Durch die Gesetzgebung und zur Erhöhung der Sicherheit wird die Benutzung von Spracherkennung bei Applikationen im Auto-
motive-Bereich in Zukunft verstärkt Anwendung finden. Neben
Telephonieanwendungen werden Sprachsteuerungen mittlerweile
auch für Telematiksysteme, Infotainment-Systeme und In-Car-
10 Systeme wie Klimaanlage eingesetzt. Das verwendete Vokabular
ist durch die aktuellen Erkennen bedingt einfach strukturiert
und in der Regel kommandobasiert.

Die Sprachsteuerung von CD-Geräten erfolgt dabei in derzeiti-
15 gen Produkten mittels Kommandos für die Grundbefehle wie etwa
"Stopp", "Play", "Pause" etc. Die Auswahl der zu spielenden
Titel wird mittels der Nummer des Titels eingegeben, also
beispielsweise durch "Play 5". Der Erkennen kann sich dabei
auf das Erkennen des Kommandowortes in Verbindung mit einer
20 Zahl beschränken. Da der Benutzer jedoch die Zuordnung der
Titel zur Nummer auf der CD oftmals nicht kennt, ist dies
Vorgehensweise unkomfortabel.

Davon ausgehend liegt der Erfindung die Aufgabe zugrunde, die
25 Bedienung von Audio- und Videogeräten einfacher, komfortabler
und sicherer zu machen.

Diese Aufgabe wird durch die in den unabhängigen Patentan-
sprüchen angegebenen Erfindungen gelöst. Vorteilhafte Ausges-
30 taltungen ergeben sich aus den abhängigen Ansprüchen.

Dementsprechend sind in einem Verfahren zur Spracherkennung
Multimediadaten auf einem Speichermedium gespeichert. Den
Multimediadaten sind Textdaten zugeordnet. In einer Graphem-
35 zu-Phonem-Konvertierung werden die Textdaten als Grapheme
Phonemen zugeordnet. Sodann können die Textdaten mit ihren

zugehörigen Phonemen als Vokabular eines Spracherkenners verwendet werden.

Dadurch ergibt sich ein sehr reduziertes und auf die jeweilige Audio- und/oder Videoanwendung spezifiziertes Erkennervokabular, das auch von einem Spracherkenner mit sehr geringen Ressourcen verarbeitet werden kann, wie er üblicherweise bei eingebetteten Spracherkennungslösungen im Auto oder in anderen Video- und/oder Audiogeräten vorliegt.

Durch diese Vorgehensweise wird es ermöglicht, einen Titel beispielsweise durch "Play Waterloo" oder nur "Waterloo" direkt einzugeben, ohne dass der Benutzer sich während der Autofahrt nebenbei noch die richtige Titelnummer überlegen muss. Speziell bei Audiosystemen mit CD-Wechslern ist ein direkter Zugriff wünschenswert.

Multimediadaten können Audio-, Video- oder Bilddaten sein. Das Speichermedium kann eine Audio-CD, eine Video-CD, eine DVD, ein mp3-Player, ein Festplatten-Videorekorder, eine Festplatte, eine Photo-CD, eine Diskette, ein USB-Stick, eine Mini-Disc oder jedes andere fest eingebaute oder wechselbare bzw. portable Speichermedium ein.

Gemäß einer Ausführungsform sind die Multimediadaten Audiodaten und das Speichermedium eine CD.

Soweit die CD CD-Text aufweist, sind die den Audiodaten zugeordneten Textdaten auf der CD als CD-Text gespeichert. Diese können dann direkt für die Graphem-zu-Phonem-Konvertierung herangezogen werden.

Die Multimediadaten können beispielsweise MP3-Daten sein. Dann sind die Textdaten vorzugsweise in einer Playlist gespeichert.

Die den Multimediadaten zugeordneten Textdaten können auch allgemein in einem Inhaltsverzeichnis des Speichermediums gespeichert sein, das die Multimediadaten enthält.

- 5 Gemäß einer Ausführungsform sind die Multimediadaten Videodaten. Das Speichermedium kann dabei zum Beispiel eine DVD sein.

- 10 Alternativ oder ergänzend können die den Multimediadaten zugeordneten Textdaten von einer zentralen Datenbank abgerufen werden, insbesondere über das Internet aus einer Internet-Datenbank.

- 15 Die Textdaten enthalten vorzugsweise den Namen des oder der Interpreten und/oder den Titel der Multimediadaten, denen sie zugeordnet sind.

- 20 Insbesondere wird über das Verfahren ein Multimediagerät mit Hilfe des Spracherkenners gesteuert. Das Multimediagerät kann ein CD-Player, ein mp3-Player, ein CD-Wechsler, ein Mini-Disc-Player, ein Videorekorder, ein DVD-Player oder ein vergleichbares Gerät sein.

- 25 In einem weiteren Schritt können die Textdaten über eine Text-zu-Sprache-Konvertierung akustisch ausgegeben werden, so dass dem Benutzer seine Wahlmöglichkeiten, insbesondere hinsichtlich Titel und Interpreten, vorgelesen werden.

- 30 Eine Anordnung, die eingerichtet ist, eines der geschilderten Verfahren auszuführen, lässt sich zum Beispiel durch Programmieren und Einrichtung einer Datenverarbeitungsanlage mit zu den genannten Verfahrensschritten gehörigen Mitteln realisieren.

- 35 Die Anordnung kann beispielsweise ein Autoradio, insbesondere integriert mit Navigationssystem, ein CD-Spieler und/oder ein DVD-Spieler sein.

Weitere Merkmale und Vorteile der Erfindung ergeben sich aus der Beschreibung von Ausführungsbeispielen.

- 5 In einem Verfahren zur Spracherkennung wird eine Graphem-zu-Phonem-Technologie bei einem eingebetteten Spracherkenner dazu verwendet, dass die Titelnamen von Songs in Phonem-Folgen konvertiert werden und als Erkennervokabular zur sprachlichen Ansteuerung von CD-, DVD- und/oder MP3-Playern eingesetzt
10 werden. Dies erlaubt dem Benutzer eine direkte Anwahl der Songs über Titel, Interpreten oder alternativ konventionell über die gewohnte Nummern-Nomenklatur.

- Werden zu den als Vokabular aufbereiteten Titeln verschiedener CDs die zugeordneten Positionen im CD-Wechsler vermerkt,
15 kann bei der sprachlichen Eingabe der Titel erkannt und einer bestimmten CD zugeordnet werden. Der Wechsler kann die gewünschte CD einlegen und den gewählten Song abspielen. Die Vokabulargröße in einem 5-fach-Wechsler mit jeweils 20 Liedern pro CD beträgt demnach ca. 100 Einträge. Dies stellt eine Vokabulargröße dar, die mit gängiger Technologie von eingebetteten Spracherkennern abgedeckt werden kann.
20

- Da Song-Titel in unterschiedlichen Sprachen vorliegen können,
25 ist vor der Konvertierung der Titel in Phonem-Folgen eine Sprachidentifikation durchzuführen, die den geeigneten Phonem-Set und die korrekten sprachspezifischen Konvertierungsregeln festlegt.

- 30 Bei Audio-CDs liegen die Song-Titel in Textform auf CD-Text-kompatiblen CDs vor. Als alternative Lösung in vernetzten Fahrzeugen kann die Titelliste per Download zur Verfügung gestellt werden.

- 35 Es werden also Textdaten von Audio- und/oder Videomedien als Vokabularbasis für den Spracherkenner verwendet. Die direkte Sprachanwahl von Songtiteln erlaubt eine komfortable und für

den Fahrer wenig ablenkende Methode zur Bedienung des CD- und MP3-Equipments in Fahrzeugen. Durch die Nutzung der Graphem-zu-Phonem-Technologie kann diese direkte Sprachanwahl realisiert werden und dem Benutzer im Rahmen seiner Sprach-
5 Bedienoberfläche zur Verfügung gestellt werden.

Das vorgestellte Verfahren ist aufgrund seiner Sichtbarkeit an der Benutzeroberfläche leicht nachweisbar. Durch die deutliche Komforterhöhung ist der Mehrwert durch den Benutzer
10 groß und erkennbar. Da sich sprecherunabhängige Systeme auf längere Frist auch im Automotive-Bereich durchsetzen werden, bietet sich eine sprachliche CD- und/oder DVD-Ansteuerung als ideale Ergänzung an.

15 Das Verfahren kann beispielsweise direkt für CDs im CD-Text-Format verwendet werden. Auf einer Audio-CD sind neben den eigentlichen Musikdaten noch Zusatzdaten gespeichert, so genannte "Sub-Channels". Es gibt dabei acht Sub-Channels (p, q, r, s, t, u, v und w). Der q-Sub-Channel enthält beispielsweise
20 Informationen über die momentane Position. Eine besondere Stellung nimmt der Leadin-Bereich ein. Der Leadin-Bereich ist ein Bereich vor den normalen Musikdaten und enthält in den q-Sub-Channels die "Table of Contents" (TOC) der CD, also das Inhaltsverzeichnis der CD. In der TOC sind die Anfangspositionen der einzelnen Tracks gespeichert. In den Sub-Channels
25 r-w des Leadins werden nun die CD-Text-Informationen gespeichert, beispielsweise der Name der CD, die Namen der Tracks sowie die Interpreten.

30 Mit dieser Information kann für den Spracherkenner ein Vokabular dynamisch erzeugt werden. Dank Graphem-zu-Phonem-Konvertierung können dabei die Textdaten in Erkenner-verständliche Phonemketten umgesetzt werden. Zur Bedienung kann dann das Vokabular oder Teile davon zur Steuerung des
35 Audio- und/oder Videogeräts verwendet werden.

Patentansprüche

1. Verfahren zur Spracherkennung,

- bei dem Multimediataten auf einem Speichermedium gespeichert sind,

- bei dem den Multimediataten jeweils Textdaten zugeordnet sind,

- bei dem Grapheme der Textdaten Phonemen zugeordnet werden,

- bei dem die Textdaten mit ihren zugehörigen Phonemen als

Vokabular eines Spracherkenners verwendet werden.

2. Verfahren nach Anspruch 1,

bei dem die Multimediataten Audiodaten sind und das Speichermedium eine CD ist.

3. Verfahren nach Anspruch 2,

bei dem die den Audiodaten zugeordneten Textdaten auf der CD als CD-Text gespeichert sind.

4. Verfahren nach einem der vorhergehenden Ansprüche,

bei dem die Multimediataten MP3-Audiodaten sind.

5. Verfahren nach Anspruch 4,

bei dem die Textdaten in einer Playlist gespeichert sind.

6. Verfahren nach Anspruch 1, bei dem die Multimediataten Videodaten sind.

7. Verfahren nach Anspruch 1, bei dem das Speichermedium eine

DVD ist.

8. Verfahren nach einem der vorhergehenden Ansprüche,

bei dem die Textdaten in einem Inhaltsverzeichnis auf dem Speichermedium gespeichert sind.

9. Verfahren nach einem der vorhergehenden Ansprüche,
bei dem die Textdaten von einer zentralen Datenbank abgerufen
werden, insbesondere über das Internet.

5 10. Verfahren nach einem der vorhergehenden Ansprüche,
bei dem die Textdaten den Namen des Interpreten und/oder den
Titel der Multimediadaten enthalten, denen sie zugeordnet
sind.

10 11. Verfahren nach einem der vorhergehenden Ansprüche,
bei dem ein Multimediagerät über den Spracherkenner gesteuert
wird.

12. Verfahren nach einem der vorhergehenden Ansprüche,
15 bei dem die Textdaten zumindest teilweise in einer Text-zu-
Sprache-Konvertierung konvertiert und akustisch ausgegeben
werden.

13. Anordnung, die eingerichtet ist, ein Verfahren nach zu-
20 mindest einem der vorstehenden Ansprüche auszuführen.

14. Anordnung nach Anspruch 13,
dadurch gekennzeichnet,
dass die Anordnung ein Auto, ein Autoradio, ein CD-Spieler
25 und/oder ein DVD-Spieler ist.